



PCT10

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/018,470A

DATE: 02/27/2003 TIME: 08:05:17

Input Set : N:\Crf4\02132003\J018470.raw
Output Set: N:\CRF4\02272003\J018470A.raw

```
1 <1110 > APPLICANT: Frazer, Claire M.
         Hickey, Erin
         Peterson, Jeremy
         Tettelin, Herve
         Venter, J. Craig
         Masignani, Vega
         Galeotti, Cesira
         Mora, Manrosa
         Ratti, Giulio
10
         Scarselli, Maria
         Scarlato, Vincenzo
11
12
         Rappuoli, Rino
         Pizza, Mariagratia
1 3
14
         Grandi, Guido
15 < 120 >  TITLE OF INVENTION: Neisseria Genomic Research
16 <1305 FILE REFERENCE: CHIR-0319
17 <140> CURRENT APPLICATION NUMBER: US/10/018,470A
18 -: 141 > CURRENT FILING DATE: 2000-03-08
19 < 150 > PRIOR APPLICATION NUMBER: 60/132,068
20 (151) PRIOR FILING DATE: 1999-04-30
21 <150% PRIOR APPLICATION NUMBER: PCT/US99/25373
33 (151) PRIOR FILING DATE: 1999-10-08
23 <150> PRIOR APPLICATION NUMBER: GB-0004695.3
24 (151) PRIOR FILING DATE: 2000-02-28
25 -150 PRIOR APPLICATION NUMBER: PCT/US/05928
He +151: PRIOR FILING DATE: 2000-03-08
\pm 7 \times 160. NUMBER OF SEQ ID NOS: \pm 107
28 - 170 - SOFTWARE: PatentIn version 3.1
30 -: 210: SEQ ID NO: 1
31 - 211. LENGTH: 2272325
32 - 212: TYPE: DNA
33 - (213) ORGANISM: Neisseria meningitidis
34 <400> SEQUENCE: 1
                                                                                   60
         taaaccttat ccacatecaa acqcataacc qtaacccatt caccqttatq qaaatqtcqc
         ccgacaacca cccagccgaa tgattcataa aatatttgca catcaggcgt ataaagatac
3 ti
                                                                                  120
                                                                                  180
37
         aagaacttta teeecaqeqa aegegetgeg eetatgeagt gggegaeeag eeteetgeea
         atgeettite egegatatte aggtaaaaca aagacateee ecaaccaata ticatacegt
                                                                                  240
3 8
34
         qqaaaacttt ecatateatq eeqettqaee qeaqeeqaac ecaacaqqat teeqqaatea
                                                                                  300
         tocacagoog caaatgecag eggeagtteg teateettea aacacetgee gtaataggea
                                                                                  360
41
         tgaatottat ocacagaaga ccacggttoa aatoogtgoo actootcaaa caacgootga
                                                                                  420
                                                                                  480
4.1
         accaacetge eqatatgeee ggettteage egtgtaatga aaacagtatt gtecacaaag
         agggaattca teggteaatt eeeegaegee ttegtteeee etgegeegta aacegeatte
                                                                                  540
4 3
```

caagcatggt ccaaacgeac tecgatttgc ctcaaatett cagectgeeg ggetttttge

600

4.1

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/018,470A

DATE: 02/27/2003 TIME: 08:05:17

Input Set : N:\Crf4\02132003\J018470.raw
Output Set: N:\CRF4\02272003\J018470A.raw

						+++	3600
94	ggcttacaca g	gggttttact	caatatcccg	cctacaaccg	taccaaacgg	tttacaatac	3660
95 •	ccgaatcgac a	atacaaagga	caaaacgatg	aaatacttga	atettgeege	aatcaccctt	3720
96	geegeeacat t	ttgccgcaca	taccgcctcg	gcagacgaac	tggccggatg	gaaagacaac	
97	accccgcaaa	geetgeaate	geteaaagee	cccgtacgca	tegteaacet	ttgggcgact	3780
98 -	tggtgcggcc	cgtgccgaaa	agagatgeet	gccatgtcca	aatggtacaa	agegeagaaa	3840
99	aaaggcagcg	tcgatatggt	eggeategeg	ctcgacacat	ccgacaatat	eggeaactte	3900
100	ctcaaacaaa	ctcctgtttc	ctacccgatt	. tggcgttaca	. eeggggegaa	cageegaaae	3960
101	tttatgaaaa	cctacggaaa	cactgtcggc	gtactgccct	. ttaccgtcgt	cgaagcaccg	4020
102	aaatgeggat	acaggcagac	cattaccggg	gaggtaaacg	aaaaaagcct	gaccgacgcc	4080
103	gtcaaactcg	cccattcaaa	atgccgttaa	acgccggatg	cegtetgaaq	, ccgcttcaga	4140
104	tggcattttt	cttttccacc	cgcctgccgg	tgcaaactta	tecaetatet	aaaaacaggc	4200
105	ggaatcttta	taatcggcac	tgtcttacct	. attgttcaga	cggcatatco	c ctgcggacgc	4260
106	aaccgcccga	aacgatatgc	cgcccttcct	. tacaggacct	cctatgatco	gtttcgaaca	4320
107	agtttccaaa	acctatcccg	gcggttttga	. agccctgaaa	aacgtcagct	tccaaatcaa	4380
108	caaaggcgaa	atgatattta	tegegggaea	. ctccggttcg	ggcaaatcca	ccatcctcaa	4440
109	actgatttcg	ggcattacca	agccgagcag	gggcaaaatc	ctgtttaacg	g ggcaggacct	4500
110	eggeaeattg	teegacaace	aaatcggctt	. tatgcgccaa	. cacateggea	ı tegtgtteca	4560
111	agaccacaaa	atcctctacg	accgcaacgt	cctgcaaaac	gtcatcctgo	c cgcttcggat	4620
112	tatoggotat	ccqccqcqca	aagccgaaga	gegtgeeege	ategecates	g aaaaagtcgg	4680
113	cctgaaagga	cgagaattgg	acgatecegt	aaccetetee	ggcggtgaad	c aacaacgcct	4740
114	gtgcatcgcc	cacaccatca	ttcaccaged	e cggcctgctg	attgccgacg	g aacceteege	4800
115	caacetegae	cgcgcctacg	cgctcgatat	. tatggaattg	r ttcaaaacct	tecaegaage	4860
116	gggaactacc	gtcatcgttg	ccgcacatga	cgaaaccctg	, atggcggact	acggacaccg	4920
117	catectgege	ctctcgaaag	gacgactcgc	atgagcatca	i tccactacct	ctcgctgcac	4980
118	gtcgaatccg	cqcqcaccqc	gctcaagcag	ctcctgcgcc	aaccetteg	g cacactgett	5040
119	acceteatga	tactcaccat	cqcqatqacc	ctgccgctgt	ttatgcatct	gggcatccaa	5100
120	agegggeaaa	acatattaga	caaactcaac	gagtegeege	aaatcacaat	ctatatggaa	5160
121	acctccqccq	cacaaaqcqa	cagegatace	gtccgcagcc	tgctggcgcg	g cgacaaacgg	5220
122	ctcgacaaca	tecgetteat	cggcaaagaa	gacggtctgg	, aagaattaca	a gtccaatctt	5280
123	gaccaaaatc	tgatttccat	gcttgacggc	aaccccctgc	cggatgtct1	tategttace	5340
134	cccaacccaa	caaccacqcc	cgcccaaatg	, caggcaatct	: accgagacat	taccaaactg	5400
135	ectatogteg	aatccgcgtc	tatggatacc	gaatgggtgc	e aaacgetgta	a ccaaatcaac	5460
136	gagttcatcc	gcaaaatttt	gtggtttctt	tecetgacge	: tggggatgg	c gttcgtcctt	5520
137	gtegeacaca	acaccatccq	cctgcaaatc	e ctcageegea	aagaagaaa1	t ogaaateace	5580
1.28	aaactcttgg	acacacccac	gtcgtttatc	e egeegeeeat	tootttato	a agccatgtgg	5640
1.29	cagageatee	tttccqccqc	egteagettg	gggetttge	gttggctgcf	t ctctgccgtg	5700
130	egeceattgg	tegatgecat	tttcaaaccc	tacggactta	a atateggete	g geggttette	5760
131	tacgetggeg	aacteggget	ggtgttcggc	ttogtcatog	, cgttgggcgf	t atteggegeg	5820
132	tagettacea	ccacccagca	cctqctcqqc	ttcaaagcca	aaaaataaa	a caccgtcaaa	5880
133	aatgeegtee	gaacccgitt	teagaeggea	tttcaattt	g ccagtataat	t ggcgcatttt	5940
134	ticcaacaagg	aacctaccat	getgaeeteg	gaacaaqtaa	aagccatgat	t tgaaggegtg	6000
135	gcaaaatgcg	aacatatcga	agtagaagge	gacggacacc	atttttcg	c egtcategtt	6060
136	teateagaat	ttgaaggcaa	ggeaegeete	gegegeeace	geetgattaa	a agacggactc	6120
137	aaagcccaac	tggaaagtaa	egaactgcac	geactiteca	a ttteggttg	e egecacteeg	6180
138	geggaatgag	cagccaaagc	acaataatco	ccacacaaaa	a atgeogtote	g aaaccatttc	6240
134	gtttcagacg	gcatttttt	tatatcaaac	cgcttacgco	g cogegitti	c ccaaagegge	6300
140	tacggcaggc	agetetttae	cttccaaqaa	ctcaaggaac	gegeegeeg	c cggtggagat	6360
141	gtagccgaft.	tgttcggtaa	egeegaattt	ggcaatege	gecagegtg	t egeegeegee	6420
142	cgcaatcgag	aacqctttqc	tttgggcaat	ggetteggea	agggettte	g taccgcctgc	6480

RAW SEQUENCE LISTING

DATE: 02/27/2003 PATENT APPLICATION: US/10/018,470A TIME: 08:05:17

Input Set : N:\Crf4\02132003\J018470.raw Output Set: N:\CRF4\02272003\J018470A.raw

	143	gaattggtca	aactcgaaca	cgccgaccgg	cccgttccaa	acgaccgtac	cggcggcttt	6540
	144	aagcaaatcg	gcaagcgcgg	cageggattt	cggaccgatg	tccaaaatca	tctcgtcttc	6600
	145	ggcaacgtcg	gcaatgtctt	tcaccacage	ttccgcatcg	gcggcaatgt	ctttcaccac	6660
	146	agetteegea	toggoggoaa	aggetttgge	aacgacgaca	teggteggea	geggeaeaga	6720
	14/	accgcctttt	gccgccattt	tegecataat	ttttttggat	tottocacca	aatcgtgttc	6780
	148	cgccaaagat	ttgccgatgg	ctttgccttc	cgccaacagg	aaggtgtttg	cgataccgcc	6840
	149	geegaegatg	agttggtcga	ctttgtccgc	cagcgattcg	aggatggtca	gcttggtgga	6900
	150	cactttgctg	ccggcaacga	tggcaaccat	egggegegeg	ggctgtttca	aggetttgee	6960
	151	caaagegteg	agttcgcccg	ccatcaatac	gccggcgcag	gcaacgggcg	cggcttgggc	7020
	152	gacggetteg	gtcgaggctt	gggcgcggtg	ggcggttccg	aacgcgtcat	tgacgaacac	7080
•	153	gtcgcacaaa	gaagcgtagg	ctttacccag	ttccaaatcg	tttttcttct	cgcctttgtt	7140
	154	gatgcgcacg	ttttgcagca	tgacgacatc	gcccgcgttc	agggcgggtt	tgttttcacg	7200
	155			egtetttgee				7260
	156			ggaactcgcc				7320
	157	gataacggac	gcaccgttgt	ccacgcagta	tttaatggac	gcgagcgagg	cgcggatacg	7380
	158	ggtgtcgtcg	ctgattttgc	cgtctttgaa	eggtaegtte	atatcggcgc	ggatgaggac	7440
	159			gttcggtcag				7500
	160	atcagtgttt	gegataegga	aacaattgat	gccgtctgaa	ggcttcagac	ggcatcgcaa	7560
	161			ctcgattttc				7620
	162			gtggtaaatc				7680
	163			ggeggaegea				7740
	164			cacaaatgcc				7800
	165			gtgcataaag				7860
	166	cagetteect	ccgccacggc	attcaatgcc	ataaactgcg	cctgcatatc	cgtggggaag	7920
	167	ccggggtgga	cgaccgtgcg	gatgtccacc	geetteggae	gctgccgcat	atcgatggcg	7980
	168	atccaatcgt	egecegeete	aatcaccgca	cctgcctcaa	ccagtttgtc	caacaccact	8040
	169	tccatcgttt	teggegegge	attccgcaaa	accaccctgc	caccggttat	egecaeegeg	8100
	170	cacaggaacg	teccegeete	gateeggteg	gggacgacgc	tgtgttcgca	gccttgcagc	8160
	171	tegtecacee	cttccacaat	cattgtggac	gtaccgatgc	cgctgatttt	cgcgcccatt	8220
	172	ttgaccaggc	attccgccaa	ategaceact	tcaggctcaa	tggcgcagtt	ttccaaaacc	8280
	173	gtcgtacctt	degecagegt	cgccgccatc	agcaggtttt	ccgtgccgcc	gacggtaacg	8340
	174	acatccatcg	ccacgcgcgt	acctttgagt	ttgcctttgg	ctttgacgta	accgtgttcg	8400
	175	ataacaatct	cagcacccat	egetteeaag	cctttcaaat	gctgatcgac	ggggcgcgaa	8460
	176	ccgatggcgc	ageegeeegg	caggetgaet	tgcgcctcgc	cgaaacgcgc	cagcgtcggg	8520
	177	cccagcacca	aaatcgaagc	gcgcatcgtt	eggaccaact	cgtaaggggc	gcaggtattg	8580
	17H	tttaccgtac	cgccgttgat	ttcaaattcg	ctgatattgt	cggtcaggac	gegegegeee	8640
	174	atcccctgaa	geagettttg	cgtggttttc	acatotycca	gcatagggac	gtttttcagg	8700
	180	cgcaacgtac	ccgatgtcag	caaacccgcg	cacatcageg	gcaatgccgc	gtttttcgcg	8760
	181	cccgagaccg	ttatttcccc	gttgagcggg	ccgtttgcgg	agattttcag	tttgtccacg	8820
	182	tttgttcttt	cctggtgggt	acttgtatag	tgaattaaca	aaaatcggga	caaggcggcg	8880
	183	aagccgcaga	cagtacagat	agtacagaac	cgattcactt	ggtgcttcag	caccttagag	8940
	184	aatcgttctc	tttgagctaa	ggcgaggcaa	taccgtactg	gtttttgtta	atccactata	9000
	185	atatttcaat	tetegggaea	acgcataaag	catcaccega	tgaaggttgc	agaggcggaa	9060
	186	ttataaggga	ttttcgggaa	aaatacggaa	geegeaceaa	agaatttgac	gaaatgeege	9120
	187	gctttccgaa	caaggattgt	cggaagacaa	aaaagccgag	ttttgaaaac	tcagcttttt	9180
	188			gagegatteg				9240
	189			gacccgataa				9300
	190			gacgggcgtt				93h0
	191	taaaccaaag	aagtcaacac	cgatgcaccc	aatgcgccga	acacgaccga	cagcgaaacg	9420

RAW SEQUENCE LISTING

DATE: 02/27/2003 PATENT APPLICATION: US/10/018,470A TIME: 08:05:17

Input Set : N:\Crf4\02132003\J018470.raw Output Set: N:\CRF4\02272003\J018470A.raw

192			atgcattacc				9480
193	aatgccaatc	catatttcag	gaagataaag	cgttccgcca	cateegeeag	caggaaatac	9540
194			tgcgaaaata				960)()
195			gatgctgtcc				9660
196	atgaccagca	ccaaaaacag	cggcgtggcg	atttttttgc	cgttttcgac	ggtaaaaaat	9720
197	ttotogoogt	gaaattccgt	geegaeegga	acgactttct	tgacggtatt	cagcagcctg	9780
198	ctgtttgcca	aatcctcttt	ctcatcgcct	togggottca	tcatgtgtat	accagtatag	4840
199	agcaggaacg	cgccaaacag	atacagaatc	cactcaaact	gctgaaccag	tgccgcgccg	9900
200	acgaaaatca	tgacggtgcg	caataccaat	gegeecaata	cgccgtacag	cagcacgcgg	9960
201	tgctgaaact	gtggtgcgac	tttgaagtag	ccgaatatca	tcaggaacac	gaaaatattg	10020
202	togactgoca	acgatttttc	caaaatgtag	ccggtaaaga	attccaatac	tttttctttt	10080
203	gegaetgeeg	cgccgtagcc	gggattgccg	gegagtteaa	aatacagcca	gcccgcgaac	10140
204			caageegete				10200
205	tggctgccgt	ttttcttcag	cgaaaacata	tccaaggcaa	tcatgaccag	cactgccgca	10260
206			eggegaeeeg				10320
207	cctgatttga	aatgtaattg	tgttaccage	tgatataaaa	categetttt	gccaaaaaga	10380
208	caatcagcag	catatgggta	aagacgacgg	cgtgtatgta	tttcgaccaa	ccgaccgtca	10440
209			acgacggcga		_		10500
210	acaggatttt	cagogtcago	atcgtaccga	aggaagtggc	aaacggttcg	cccaatatag	10560
211			atcacgatgo				10620
212			taagacattg				10680
213	-		acqqacaaaa				10740
214	aggcaatagc	qcaatacaqa	tgaacgatgt	gegegaegge	ataaatactc	atacqatqct	10800
215		-	cggattgtat				10860
216			totogogoco		-	-	10920
217			cecececee				10980
218	, , ,	, ,	gacttgaacc				11040
219			tgaaaagccg				11100
220			cgaatcgacc				11160
221			ctcgcccata				11220
322	, ,	-	gegeegetga		-	-	11280
223			gggatggege				11340
224	-		cagtttttcc				11400
225			coggootoog	•	-		11460
226			gatgtcgtat				11520
127			egeegeatee				11580
228		-	caggoggegg	-			11640
229			aaacagcggc				11700
230		_	catcaaaatq				11760
J31	2 2 3		gtcatacqqa	-			11820
232		-	cgtcggggtc	, , , , ,			11880
233			tatcaaggac				11940
234			ggtaacggga				12000
235			ggcacttccg				12060
23h			tacatcataa				12120
1:37			gttcagacgg				12180
238			ccettgeecg				12240
239			gtgtgcgcgc				12300
240			gtgtageegt				12360

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/10/018,470A

TIME: 02/27/2003
TIME: 08:05:18

Input Set : N:\Crf4\02132003\J018470.raw
Output Set: N:\CRF4\02272003\J018470A.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the $\langle 220 \rangle$ to $\langle 223 \rangle$ fields of each sequence which presents at least one n or Xaa.

Seq#:6; N Pos. 6,8,38,39,96,102,341,390,410 Seq#:7; Xaa Pos. 3,13,114,137 Seq#:40; N Pos. 364,365,366,367,368,369,370,371,372,373,374,375,376,377,378 Seq#:40; N Pos. 379,380,381,382,383,384,385,386,387,388,389,390,391,392,393 Seq#:40; N Pos. 394,395,396,397,398,399,400,401,402,403,404,405,406,407,408 Seq#:40; N Pos. 409,410,411,412,413,414,415,416,417,418,419,420,421,422,423 Seq#:40; N Pos. 424,425,426,427,428,429,430,431,432,433,434,435,436,437,438 Seq#:40; N Pos. 439,440,441,442,443,444,445,446,447,448,449,450,451,452,453 Seq#:40; N Pos. 454,455,456,457,458,459,460,461,462,463,464,465,466,467,468 Seq#:40; N Pos. 469,470,471,472,473,474,475,476,477,478,479,480,481,482,483 Seq#:40; N Pos. 484,485,486,487,488,489,490,491,492,493,494,495,496,497,498 Seq#:40; N Pos. 499,500,501,502,503,504,505,506,895,982 Seq#:41; Xaa Pos. 122,123,124,125,126,127,128,129,130,131,132,133,134,135 Seq#:41; Xaa Pos. 136,137,138,139,140,141,142,143,144,145,146,147,148,149 Seq#:41; Xaa Pos. 150,151,152,153,154,155,156,157,158,159,160,161,162,163 Seq#:41; Xaa Pos. 164,165,166,167,168,169,170,171,172,173,174,175,176,177 Seq#:41; Xaa Pos. 178,179,180,181,182,183,184,185,186,187,188,189,190,191 Seq#:41; Xaa Pos. 192,193,194,195,196,197,198,199,200,201,202,332,361 Seq#:46; N Pos. 995,1082 Seq#:47; Xaa Pos. 332,361 Seq#:50; N Pos. 1216,1334,1347 Seq#:51; Xaa Pos. 149,158,167,190,198,289,313,328,331,339,343,347,406,445 Seq#:63; Xaa Pos. 116

VERIFICATION SUMMARY

DATE: 02/27/2003 TIME: ()8:05:18 PATENT APPLICATION: US/10/018,470A

Input Set : N:\Crf4\02132003\J018470.raw Output Set: N:\CRF4\02272003\J018470A.raw

L:38026 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6 after pos.:0 M:341 Repeated in SeqNo=6 $L:38056\ M:341\ W:$ (46) "n" or "Xaa" used, for SEQ ID#:7 after pos.:0 M:341 Repeated in SeqNo=7 L:39195.M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:40 after pos.:360 M:341 Repeated in SeqNo=40 L:39234 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:41 after pos.:112 M:341 Repeated in SeqNo=41 L:39450 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:46 after pos.:960 M:341 Repeated in SeqNo=46 L:39507 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:47 after pos.:320 M:341 Repeated in SeqNo=47 L:39628 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:50 after pos.:1200 M:341 Repeated in SeqNo=50 L:39712 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:51 after pos.:144 M:341 Repeated in SeqNo=51 L:40420 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:63 after pos.:112

US 1001847008P1



Creation date: 10-15-2003

Indexing Officer: ATEKLY - ALEM TEKLAY

Team: OIPEScanning Dossier: 10018470

Legal Date: 03-05-2003

No.	Doccode	Number of pages
1	M903	2

Total number of pages: 2

Remarks:

Order of re-scan issued on